

Reply to Office Action of June 16, 2004

AMENDMENTS TO THE SPECIFICATION

Please replace the last paragraph on page 1 and going onto page 2 with the following amended paragraph:

The invention thus proposes resolving this problem in a manner that is simple and of low cost by implementing a method of recognizing and indexing documents using a scanner connected to a computer, the method including scanning the documents, then using a pointing device or member of the computer to designate an arbitrary point P in at least one box of the documents, and finally recognizing by OCR the characters in ~~saïd~~ the box so as to store them in a first database connected to the computer to enable documents scanned in this way to be indexed. The designation step comprises a step of searching for and identifying the box of the document which contains ~~saïd~~ the point P designated by the user.

Please replace the third paragraph on page 2 with the following amended paragraph:

The step of looking for and identifying ~~saïd~~ the box can be performed by applying a shape search algorithm over a determined zone surrounding ~~saïd~~ the point P as previously designated by the user. This shape search algorithm can be an algorithm based on the Hough transform, or a projection algorithm which counts the number of pixels present in each vertical or horizontal line of ~~saïd~~ the determined search zone and, on the basis of those numbers, examines peaks in X and Y projection profiles to find the horizontal and vertical lines present in the search zone.

Please replace the fifth paragraph on page 2 with the following amended paragraph:

The scanning step is preferably performed initially for all of the documents to be processed, with ~~saïd~~ the steps of identifying the box and of performing OCR on its content being performed subsequently in succession for each of the documents. Nevertheless, the scanning step could equally well be performed initially for a first document, with ~~saïd~~ the steps of identifying

Reply to Office Action of June 16, 2004

the box and performing OCR on its content then being performed for the same document, and with these three steps then being repeated successively for all of the documents to be processed.

Please replace the title on line 4 of page 4 with the following amended title:

Detailed description of preferred embodiments.

Please replace the last paragraph on page 5 and going onto page 6 with the following amended paragraph:

The method implemented in the above apparatus is illustrated in Figure 4 and thus follows the following steps. After a first document has been scanned in a first step 200 by the scanner 30 associated with the software 100, the image of the document is stored temporarily in a second step 210 in the memory of the computer 32 and possibly simultaneously it is displayed on the display screen of the computer (when necessary after performing a magnification or "zoom" operation). If the processing software 130 cannot identify the type of document that has been scanned on the basis of the data from the characterization database 38 (test in step 220), then this identification is performed during the following steps of the process, and in particular a step 230 is performed initially using the pointing member 36 associated with its software 130 whereby the user designates a point P in a first predetermined zone of the document, e.g. the box 16 in the information block 14 of the drawing that contains the identification number of the drawing. Optionally, and as shown in dashed lines in step 240, the user can then specify the type of character to be recognized in this box. This indication can serve to restrict the choice of characters to be recognized (e.g. only the digits 0 to 9), thereby improving the subsequent OCR step. Once this point has been designated (with its coordinates then being determined relative to a predetermined origin point), a new step 250 is performed in which the box containing the point P is searched for and identified (i.e. the boundary lines of the box are searched for and identified as explained in greater detail with reference to Figure 5), and once the box has been identified (e.g. with the outline of the box being displayed at increased brightness or in another color) its characterization elements are stored in the second database 38 in a step 260 (the dimensions of the box and the position of its center are also stored), and in the immediately